```
LOCUS       HUMHDA                  10348 bp    mRNA    linear   PRI 13-MAY-2002
DEFINITION  Homo sapiens Huntington's Disease (HD) mRNA, complete cds.
ACCESSION   L12392
VERSION     L12392.1  GI:1709991
KEYWORDS    trinucleotide repeat.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 10348)
  AUTHORS   MacDonald,M. and Ambrose,C.M.
  TITLE     A novel gene containing a trinucleotide repeat that is expanded and
            unstable on Huntington's disease chromosomes. The Huntington's
            Disease Collaborative Research Group [see comments]
  JOURNAL   Cell 72 (6), 971-983 (1993)
   PUBMED   8458085
COMMENT     On Dec 9, 1996 this sequence version replaced gi:454414.
FEATURES             Location/Qualifiers
     source          1..10348
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /note="derived from DNA and mRNA sequence"
     gene            1..10348
                     /gene="HD"
     CDS             316..9750
                     /gene="HD"
                     /codon_start=1
                     /product="Huntington's Disease protein"
                     /protein_id="AAB38240.1"
                     /db_xref="GI:454415"
                     /db_xref="GDB:G00-119-307"
                     /translation="MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQQQQPPPP
                     PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPGPAVAEEPLHRPKKELSATKKDRV
                     NHCLTICENIVAQSVRNSPEFQKLLGIAMELFLLCSDDAESDVRMVADECLNKVIKAL
                     MDSNLPRLQLELYKEIKKNGAPRSLRAALWRFAELAHLVRPQKCRPYLVNLLPCLTRT
                     SKRPEESVQETLAAAVPKIMASFGNFANDNEIKVLLKAFIANLKSSSPTIRRTAAGSA
                     VSICQHSRRTQYFYSWLLNVLLGLLVPVEDEHSTLLILGVLLTLRYLVPLLQQQVKDT
                     SLKGSFGVTRKEMEVSPSAEQLVQVYELTLHHTQHQDHNVVTGALELLQQLFRTPPPE
                     LLQTLTAVGGIGQLTAAKEESGGRSRSGSIVELIAGGGSSCSPVLSRKQKGKVLLGEE
                     EALEDDSESRSDVSSSALTASVKDEISGELAASSGVSTPGSAGHDIITEQPRSQHTLQ
                     ADSVDLASCDLTSSATDGDEEDILSHSSSQVSAVPSDPAMDLNDGTQASSPISDSSQT
                     TTEGPDSAVTPSDSSEIVLDGTDNQYLGLQIGQPQDEDEEATGILPDEASEAFRNSSM
                     ALQQAHLLKNMSHCRQPSDSSVDKFVLRDEATEPGDQENKPCRIKGDIGQSTDDDSAP
```

```
                     LVHCVRLLSASFLLTGGKNVLVPDRDVRVSVKALALSCVGAAVALHPESFFSKLYKVP
                     LDTTEYPEEQYVSDILNYIDHGDPQVRGATAILCGTLICSILSRSRFHVGDWMGTIRT
                     LTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVL
                     TLRNSSYWLVRTELLETLAEIDFRLVSFLEAKAENLHRGAHHYTGLLKLQERVLNNVV
                     IHLLGDEDPRVRHVAAASLIRLVPKLFYKCDQGQADPVVAVARDQSSVYLKLLMHETQ
                     PPSHFSVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTTRALTFGCCEA
                     LCLLSTAFPVCIWSLGWHCGVPPLSASDESRKSCTVGMATMILTLLSSAWFPLDLSAH
                     QDALILAGNLLAASAPKSLRSSWASEEEANPAATKQEEVWPALGDRALVPMVEQLFSH
                     LLKVINICAHVLDDVAPGPAIKAALPSLTNPPSLSPIRRKGKEKEPGEQASVPLSPKK
                     GSEASAASRQSDTSGPVTTSKSSSLGSFYHLPSYLKLHDVLKATHANYKVTLDLQNST
                     EKFGGFLRSALDVLSQILELATLQDIGKCVEEILGYLKSCFSREPMMATVCVQQLLKT
                     LFGTNLASQFDGLSSNPSKSQGRAQRLGSSSVRPGLYHYCFMAPYTHFTQALADASLR
                     NMVQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHIRLFEPLVIKALK
                     QYTTTTCVQLQKQVLDLLAQLVQLRVNYCLLDSDQVFIGFVLKQFEYIEVGQFRESEA
                     IIPNIFFFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKAVTHAIPALQPIVHDLF
                     VLRGTNKADAGKELETQKEVVVSMLLRLIQYHQVLEMFILVLQQCHKENEDKWKRLSR
                     QIADIILPMLAKQQMHIDSHEALGVLNTLFEILAPSSLRPVDMLLRSMFVTPNTMASV
                     STVQLWISGILAILRVLISQSTEDIVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
                     EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVTKQLKVEMSEQQHTFYCQELGTLLMC
                     LIHIFKSGMFRRITAAATRLFRSDGCGGSFYTLDSLNLRARSMITTHPALVLLWCQIL
                     LLVNHTDYRWWAEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLAAKLGMCNREIVRRGA
                     LILFCDYVCQNLHDSEHLTWLIVNHIQDLISLSHEPPVQDFISAVHRNSAASGLFIQA
                     IQSRCENLSTPTMLKKTLQCLEGIHLSQSGAVLTLYVDRLLCTPFRVLARMVDILACR
                     RVEMLLAANLQSSMAQLPMEELNRIQEYLQSSGLAQRHQRLYSLLDRFRLSTMQDSLS
                     PSPPVSSHPLDGDGHVSLETVSPDKDWYVHLVKSQCWTRSDSALLEGAELVNRIPAED
                     MNAFMMNSEFNLSLLAPCLSLGMSEISGGQKSALFEAAREVTLARVSGTVQQLPAVHH
                     VFQPELPAEPAAYWSKLNDLFGDAALYQSLPTLARALAQYLVVVSKLPSHLHLPPEKE
                     KDIVKFVVATLEALSWHLIHEQIPLSLDLQAGLDCCCLALQLPGLWSVVSSTEFVTHA
                     CSLIYCVHFILEAVAVQPGEQLLSPERRTNTPKAISEEEEEVDPNTQNPKYITAACEM
                     VAEMVESLQSVLALGHKRNSGVPAFLTPLLRNIIISLARLPLVNSYTRVPPLVWKLGW
                     SPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYRINTLGWTSRTQFEETWATLLGVLVTQ
                     PLVMEQEESPPEEDTERTQINVLAVQAITSLVLSAMTVPVAGNPAVSCLEQQPRNKPL
                     KALDTRFGRKLSIIRGIVEQEIQAMVSKRENIATHHLYQAWDPVPSLSPATTGALISH
                     EKLLLQINPERELGSMSYKLGQVSIHSVWLGNSITPLREEEWDEEEEEEADAPAPSSP
                     PTSPVNSRKHRAGVDIHSCSQFLLELYSRWILPSSSARRTPAILISEVVRSLLVVSDL
                     FTERNQFELMYVTLTELRRVHPSEDEILAQYLVPATCKAAAVLGMDKAVAEPVSRLLE
                     STLRSSHLPSRVGALHGVLYVLECDLLDDTAKQLIPVISDYLLSNLKGIAHCVNIHSQ
                     QHVLVMCATAFYLIENYPLDVGPEFSASIIQMCGVMLSGSEESTPSIIYHCALRGLER
                     LLLSEQLSRLDAESLVKLSVDRVNVHSPHRAMAALGLMLTCMYTGKEKVSPGRTSDPN
                     PAAPDSESVIVAMERVSVLFDRIRKGFPCEARVVARILPQFLDDFFPPQDIMNKVIGE
                     FLSNQQPYPQFMATVVYKVFQTLHSTGQSSMVRDWVMLSLSNFTQRAPVAMATWSLSC
                     FFVSASTSPWVAAILPHVISRMGKLEQVDVNLFCLVATDFYRHQIEEELDRRAFQSVL
                     EVVAAPGSPYHRLLTCLRNVHKVTTC"
     polyA_site      10348
                     /gene="HD"
ORIGIN
        1 ttgctgtgtg aggcagaacc tgcgggggca ggggcgggct ggttccctgg ccagccattg
       61 gcagagtccg caggctaggg ctgtcaatca tgctggccgg cgtggccccg cctccgccgg
      121 cgcggccccg cctccgccgg cgcacgtctg ggacgcaagg cgccgtgggg gctgccggga
      181 cgggtccaag atggacggcc gctcaggttc tgcttttacc tgcggcccag agccccattc
      241 attgccccgg tgctgagcgg cgccgcgagt cggcccgagg cctccgggga ctgccgtgcc
      301 gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccttcga gtccctcaag
      361 tccttccagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag
      421 cagcagcagc aacagccgcc accgccgccg ccgccgccgc cgcctcctca gcttcctcag
      481 ccgccgccgc aggcacagcc gctgctgcct cagccgcagc cgcccccgcc gccgcccccg
      541 ccgccacccg gcccggctgt ggctgaggag ccgctgcacc gaccaaagaa agaactttca
      601 gctaccaaga aagaccgtgt gaatcattgt ctgacaatat gtgaaaacat agtggcacag
      661 tctgtcagaa attctccaga atttcagaaa cttctgggca tcgctatgga acttttctg
      721 ctgtgcagtg atgacgcaga gtcagatgtc aggatggtgg ctgacgaatg cctcaacaaa
      781 gttatcaaag ctttgatgga ttctaatctt ccaaggttac agctcgagct ctataaggaa
```

```
 841 attaaaaaga atggtgcccc tcggagtttg cgtgctgccc tgtggaggtt tgctgagctg
 901 gctcacctgg ttcggcctca gaaatgcagg ccttacctgg tgaaccttct gccgtgcctg
 961 actcgaacaa gcaagagacc cgaagaatca gtccaggaga ccttggctgc agctgttccc
1021 aaaattatgg cttcttttgg caattttgca aatgacaatg aaattaaggt tttgttaaag
1081 gccttcatag cgaacctgaa gtcaagctcc cccaccattc ggcggacagc ggctggatca
1141 gcagtgagca tctgccagca ctcaagaagg acacaatatt tctatagttg gctactaaat
1201 gtgctcttag gcttactcgt tcctgtcgag gatgaacact ccactctgct gattcttggc
1261 gtgctgctca ccctgaggta tttggtgccc ttgctgcagc agcaggtcaa ggacacaagc
1321 ctgaaaggca gcttcggagt gacaaggaaa gaaatggaag tctctccttc tgcagagcag
1381 cttgtccagg tttatgaact gacgttacat catacacagc accaagacca caatgttgtg
1441 accggagccc tggagctgtt gcagcagctc ttcagaacgc ctccacccga gcttctgcaa
1501 accctgaccg cagtcggggg cattgggcag ctcaccgctg ctaaggagga gtctggtggc
1561 cgaagccgta gtgggagtat tgtggaactt atagctggag ggggttcctc atgcagccct
1621 gtcctttcaa gaaaacaaaa aggcaaagtg ctcttaggag aagaagaagc cttggaggat
1681 gactctgaat cgagatcgga tgtcagcagc tctgccttaa cagcctcagt gaaggatgag
1741 atcagtggag agctggctgc ttcttcaggg gtttccactc cagggtcagc aggtcatgac
1801 atcatcacag aacagccacg gtcacagcac acactgcagg cggactcagt ggatctggcc
1861 agctgtgact tgacaagctc tgccactgat ggggatgagg aggatatctt gagccacagc
1921 tccagccagg tcagcgccgt cccatctgac cctgccatgg acctgaatga tgggacccag
1981 gcctcgtcgc ccatcagcga cagctcccag accaccaccg aagggcctga ttcagctgtt
2041 acccccttcag acagttctga aattgtgtta gacggtaccg acaaccagta tttgggcctg
2101 cagattggac agccccagga tgaagatgag gaagccacag gtattcttcc tgatgaagcc
2161 tcggaggcct tcaggaactc ttccatggcc cttcaacagg cacatttatt gaaaaacatg
2221 agtcactgca ggcagccttc tgacagcagt gttgataaat ttgtgttgag agatgaagct
2281 actgaaccgg gtgatcaaga aaacaagcct tgccgcatca aaggtgacat tggacagtcc
2341 actgatgatg actctgcacc tcttgtccat tgtgtccgcc ttttatctgc ttcgtttttg
2401 ctaacagggg gaaaaaatgt gctggttccg gacagggatg tgagggtcag cgtgaaggcc
2461 ctggccctca gctgtgtggg agcagctgtg gccctccacc cggaatcttt cttcagcaaa
2521 ctctataaag ttcctcttga caccacggaa taccctgagg aacagtatgt ctcagacatc
2581 ttgaactaca tcgatcatgg agacccacag gttcgaggag ccactgccat tctctgtggg
2641 accctcatct gctccatcct cagcaggtcc cgcttccacg tgggagattg gatgggcacc
2701 attagaaccc tcacaggaaa tacattttct ttggcggatt gcattccttt gctgcggaaa
2761 acactgaagg atgagtcttc tgttacttgc aagttagctt gtacagctgt gaggaactgt
2821 gtcatgagtc tctgcagcag cagctacagt gagttaggac tgcagctgat catcgatgtg
2881 ctgactctga ggaacagttc ctattggctg gtgaggacag agcttctgga aacccttgca
2941 gagattgact tcaggctggt gagctttttg gaggcaaaag cagaaaactt acacagaggg
3001 gctcatcatt atacagggct tttaaaactg caagaacgag tgctcaataa tgttgtcatc
3061 catttgcttg gagatgaaga ccccagggtg cgacatgttg ccgcagcatc actaattagg
3121 cttgtcccaa agctgtttta taaatgtgac caaggacaag ctgatccagt agtggccgtg
3181 gcaagagatc aaagcagtgt ttacctgaaa cttctcatgc atgagacgca gcctccatct
3241 catttctccg tcagcacaat aaccagaata tatagaggct ataacctact accaagcata
3301 acagacgtca ctatggaaaa taacctttca agagttattg cagcagtttc tcatgaacta
3361 atcacatcaa ccaccagagc actcacattt ggatgctgtg aagctttgtg tcttctttcc
3421 actgccttcc cagtttgcat ttggagttta ggttggcact gtggagtgcc tccactgagt
3481 gcctcagatg agtctaggaa gagctgtacc gttgggatgg ccacaatgat tctgaccctg
3541 ctctcgtcag cttggttccc attggatctc tcagcccatc aagatgcttt gattttggcc
3601 ggaaacttgc ttgcagccag tgctcccaaa tctctgagaa gttcatgggc ctctgaagaa
3661 gaagccaacc cagcagccac caagcaagag gaggtctggc cagccctggg ggaccgggcc
3721 ctggtgccca tggtggagca gctcttctct cacctgctga aggtgattaa catttgtgcc
3781 cacgtcctgg atgacgtggc tcctggaccc gcaataaagg cagccttgcc ttctctaaca
3841 aacccccctt ctctaagtcc catccgacga aaggggaagg agaaagaacc aggagaacaa
3901 gcatctgtac cgttgagtcc caagaaaggc agtgaggcca gtgcagcttc tagacaatct
3961 gatacctcag gtcctgttac aacaagtaaa tcctcatcac tggggagttt ctatcatctt
4021 ccttcatacc tcaaactgca tgatgtcctg aaagctacac acgctaacta caaggtcacg
4081 ctggatcttc agaacagcac ggaaaagttt ggagggtttc tccgctcagc cttggatgtt
4141 ctttctcaga tactagagct ggccacactg caggacattg ggaagtgtgt tgaagagatc
4201 ctaggatacc tgaaatcctg ctttagtcga gaaccaatga tggcaactgt ttgtgttcaa
4261 caattgttga agactctctt tggcacaaac ttggcctccc agtttgatgg cttatcttcc
4321 aaccccagca agtcacaagg ccgagcacag cgccttggct cctccagtgt gaggccaggc
4381 ttgtaccact actgcttcat ggcccccgtac acccacttca cccaggccct cgctgacgcc
```

```
4441 agcctgagga acatggtgca ggcggagcag gagaacgaca cctcgggatg gtttgatgtc
4501 ctccagaaag tgtctaccca gttgaagaca aacctcacga gtgtcacaaa gaaccgtgca
4561 gataagaatg ctattcataa tcacattcgt ttgtttgaac ctcttgttat aaaagcttta
4621 aaacagtaca cgactacaac atgtgtgcag ttacagaagc aggttttaga tttgctggcg
4681 cagctggttc agttacgggt taattactgt cttctggatt cagatcaggt gtttattggc
4741 tttgtattga aacagtttga atacattgaa gtgggccagt tcagggaatc agaggcaatc
4801 attccaaaca tctttttctt cttggtatta ctatcttatg aacgctatca ttcaaaacag
4861 atcattggaa ttcctaaaat cattcagctc tgtgatggca tcatggccag tggaaggaag
4921 gctgtgacac atgccatacc ggctctgcag cccatagtcc acgacctctt tgtattaaga
4981 ggaacaaata aagctgatgc aggaaaagag cttgaaaccc aaaaagaggt ggtggtgtca
5041 atgttactga gactcatcca gtaccatcag gtgttggaga tgttcattct tgtcctgcag
5101 cagtgccaca aggagaatga agacaagtgg aagcgactgt ctcgacagat agctgacatc
5161 atcctcccaa tgttagccaa acagcagatg cacattgact ctcatgaagc ccttggagtg
5221 ttaaatacat tatttgagat tttggcccct tcctccctcc gtccggtaga catgctttta
5281 cggagtatgt tcgtcactcc aaacacaatg gcgtccgtga gcactgttca actgtggata
5341 tcgggaattc tggccatttt gagggttctg atttcccagt caactgaaga tattgttctt
5401 tctcgtattc aggagctctc cttctctccg tatttaatct cctgtacagt aattaatagg
5461 ttaagagatg gggacagtac ttcaacgcta gaagaacaca gtgaagggaa acaaataaag
5521 aatttgccag aagaaacatt ttcaaggttt ctattacaac tggttggtat tcttttagaa
5581 gacattgtta caaaacagct gaaggtggaa atgagtgagc agcaacatac tttctattgc
5641 caggaactag gcacactgct aatgtgtctg atccacatct tcaagtctgg aatgttccgg
5701 agaatcacag cagctgccac taggctgttc cgcagtgatg gctgtggcgg cagtttctac
5761 accctggaca gcttgaactt gcgggctcgt tccatgatca ccacccaccc ggccctggtg
5821 ctgctctggt gtcagatact gctgcttgtc aaccacaccg actaccgctg gtgggcagaa
5881 gtgcagcaga ccccgaaaag acacagtctg tccagcacaa agttacttag tccccagatg
5941 tctggagaag aggaggattc tgacttggca gccaaacttg gaatgtgcaa tagagaaata
6001 gtacgaagag gggctctcat tctcttctgt gattatgtct gtcagaacct ccatgactcc
6061 gagcacttaa cgtggctcat tgtaaatcac attcaagatc tgatcagcct ttcccacgag
6121 cctccagtac aggacttcat cagtgccgtt catcggaact ctgctgccag cggcctgttc
6181 atccaggcaa ttcagtctcg ttgtgaaaac ctttcaactc caaccatgct gaagaaaact
6241 cttcagtgct tggaggggat ccatctcagc cagtcgggag ctgtgctcac gctgtatgtg
6301 gacaggcttc tgtgcacccc tttccgtgtg ctggctcgca tggtcgacat ccttgcttgt
6361 cgccgggtag aaatgcttct ggctgcaaat ttacagagca gcatggccca gttgccaatg
6421 gaagaactca acagaatcca ggaatacctt cagagcagcg ggctcgctca gagacaccaa
6481 aggctctatt ccctgctgga caggtttcgt ctctccacca tgcaagactc acttagtccc
6541 tctcctccag tctcttccca cccgctggac ggggatgggc acgtgtcact ggaaacagtg
6601 agtccggaca aagactggta cgttcatctt gtcaaatccc agtgttggac caggtcagat
6661 tctgcactgc tggaaggtgc agagctggtg aatcggattc ctgctgaaga tatgaatgcc
6721 ttcatgatga actcggagtt caacctaagc ctgctagctc catgcttaag cctagggatg
6781 agtgaaattt ctggtggcca gaagagtgcc ctttttgaag cagcccgtga ggtgactctg
6841 gcccgtgtga gcggcaccgt gcagcagctc cctgctgtcc atcatgtctt ccagcccgag
6901 ctgcctgcag agccggcggc ctactggagc aagttgaatg atctgtttgg ggatgctgca
6961 ctgtatcagt ccctgcccac tctggcccgg gccctggcac agtacctggt ggtggtctcc
7021 aaactgccca gtcatttgca ccttcctcct gagaaagaga aggacattgt gaaattcgtg
7081 gtggcaaccc ttgaggccct gtcctggcat ttgatccatg agcagatccc gctgagtctg
7141 gatctccagg cagggctgga ctgctgctgc ctggccctgc agctgcctgg cctctggagc
7201 gtggtctcct ccacagagtt tgtgacccac gcctgctccc tcatctactg tgtgcacttc
7261 atcctggagg ccgttgcagt gcagcctgga gagcagcttc ttagtccaga aagaaggaca
7321 aataccccaa aagccatcag cgaggaggag gaggaagtag atccaaacac acagaatcct
7381 aagtatatca ctgcagcctg tgagatggtg gcagaaatgg tggagtctct gcagtcggtg
7441 ttggccttgg gtcataaaag gaatagcggc gtgccggcgt ttctcacgcc attgctcagg
7501 aacatcatca tcagcctggc ccgcctgccc cttgtcaaca gctacacacg tgtgccccca
7561 ctggtgtgga agcttggatg gtcacccaaa ccgggagggg attttggcac agcattccct
7621 gagatccccg tggagttcct ccaggaaaag gaagtcttta aggagttcat ctaccgcatc
7681 aacacactag gctggaccag tcgtactcag tttgaagaaa cttgggccac cctccttggt
7741 gtcctggtga cgcagcccct cgtgatggag caggaggaga gcccaccaga agaagacaca
7801 gagaggaccc agatcaacgt cctggccgtg caggccatca cctcactggt gctcagtgca
7861 atgactgtgc ctgtggccgg caacccagct gtaagctgct tggagcagca gccccggaac
7921 aagcctctga aagctctcga caccaggttt gggaggaagc tgagcattat cagagggatt
7981 gtggagcaag agattcaagc aatggtttca aagagagaga atattgccac ccatcattta
```

```
     8041 tatcaggcat gggatcctgt cccttctctg tctccggcta ctacaggtgc cctcatcagc
     8101 cacgagaagc tgctgctaca gatcaacccc gagcgggagc tggggagcat gagctacaaa
     8161 ctcggccagg tgtccataca ctccgtgtgg ctggggaaca gcatcacacc cctgagggag
     8221 gaggaatggg acgaggaaga ggaggaggag gccgacgccc ctgcaccttc gtcaccaccc
     8281 acgtctccag tcaactccag gaaacaccgg gctggagttg acatccactc ctgttcgcag
     8341 tttttgcttg agttgtacag ccgctggatc ctgccgtcca gctcagccag gaggaccccg
     8401 gccatcctga tcagtgaggt ggtcagatcc cttctagtgg tctcagactt gttcaccgag
     8461 cgcaaccagt ttgagctgat gtatgtgacg ctgacagaac tgcgaagggt gcacccttca
     8521 gaagacgaga tcctcgctca gtacctggtg cctgccacct gcaaggcagc tgccgtcctt
     8581 gggatggaca aggccgtggc ggagcctgtc agccgcctgc tggagagcac gctcaggagc
     8641 agccacctgc ccagcagggt tggagccctg cacggcgtcc tctatgtgct ggagtgcgac
     8701 ctgctggacg acactgccaa gcagctcatc ccggtcatca gcgactatct cctctccaac
     8761 ctgaaaggga tcgcccactg cgtgaacatt cacagccagc agcacgtact ggtcatgtgt
     8821 gccactgcgt tttacctcat tgagaactat cctctggacg tagggccgga attttcagca
     8881 tcaataatac agatgtgtgg ggtgatgctg tctggaagtg aggagtccac cccctccatc
     8941 atttaccact gtgccctcag aggcctggag cgcctcctgc tctctgagca gctctcccgc
     9001 ctggatgcag aatcgctggt caagctgagt gtggacagag tgaacgtgca cagcccgcac
     9061 cgggccatgg cggctctggg cctgatgctc acctgcatgt acacaggaaa ggagaaagtc
     9121 agtccgggta gaacttcaga ccctaatcct gcagcccccg acagcgagtc agtgattgtt
     9181 gctatggagc gggtatctgt tctttttgat aggatcagga aaggctttcc ttgtgaagcc
     9241 agagtggtgg ccaggatcct gccccagttt ctagacgact tcttcccacc ccaggacatc
     9301 atgaacaaag tcatcggaga gtttctgtcc aaccagcagc catacccccа gttcatggcc
     9361 accgtggtgt ataaggtgtt tcagactctg cacagcaccg ggcagtcgtc catggtccgg
     9421 gactgggtca tgctgtccct ctccaacttc acgcagaggg ccccggtcgc catggccacg
     9481 tggagcctct cctgcttctt tgtcagcgcg tccaccagcc cgtgggtcgc ggcgatcctc
     9541 ccacatgtca tcagcaggat gggcaagctg gagcaggtgg acgtgaacct tttctgcctg
     9601 gtcgccacag acttctacag acaccagata gaggaggagc tcgaccgcag ggccttccag
     9661 tctgtgcttg aggtggttgc agccccagga agcccatatc accggctgct gacttgttta
     9721 cgaaatgtcc acaaggtcac cacctgctga gcgccatggt gggagagact gtgaggcggc
     9781 agctggggcc ggagcctttg gaagtctgtg cccttgtgcc ctgcctccac cgagccagct
     9841 tggtccctat gggcttccgc acatgccgcg ggcggccagg caacgtgcgt gtctctgcca
     9901 tgtggcagaa gtgctctttg tggcagtggc caggcaggga gtgtctgcag tcctggtggg
     9961 gctgagcctg aggccttcca gaaagcagga gcagctgtgc tgcaccccat gtgggtgacc
    10021 aggtcctttc tcctgatagt cacctgctgg ttgttgccag gttgcagctg ctcttgcatc
    10081 tgggccagaa gtcctccctc ctgcaggctg gctgttggcc cctctgctgt cctgcagtag
    10141 aaggtgccgt gagcaggctt tgggaacact ggcctgggtc tccctggtgg ggtgtgcatg
    10201 ccacgccccg tgtctggatg cacagatgcc atggcctgtg ctgggccagt ggctgggggt
    10261 gctagacacc cggcaccatt ctcccttctc tcttttcttc tcaggattta aaatttaatt
    10321 atatcagtaa agagattaat tttaacgt
//
```